A maximum principle for forward-backward stochastic Volterra integral equations and applications in finance*

Tianxiao Wang and Yufeng Shi †
School of Mathematics, Shandong University, Jinan 250100, China

April 12 2010

Abstract

This paper formulates and studies a stochastic maximum principle for forward-backward stochastic Volterra integral equations (FBSVIEs in short), while the control area is assumed to be convex. Then a linear quadratic (LQ in short) problem for backward stochastic Volterra integral equations (BSVIEs in short) is present to illustrate the aforementioned optimal control problem. Motivated by the technical skills in solving above problem, a more convenient and briefer method for the unique solvability of M-solution for BSVIEs is proposed. At last, we will investigate a risk minimization problem by means of the maximum principle for FBSVIEs. Closed-form optimal portfolio is obtained in some special cases.

Keywords: Forward-backward stochastic Volterra integral equations, Adapted M-solution, Optimal control, Stochastic maximum principle, Backward linear quadratic, Risk minimization problem

1 Introduction

Throughout this paper we assume that all uncertainties come from a common complete probability space $(\Omega, \mathcal{F}, P)$ on which is defined a d-dimensional Wiener process $(W_t)_{t\in[0,T]}$. The main objective of this paper is to study the optimal control problem for the following forward-backward stochastic Volterra integral equation (FBSVIE in short)

$$\begin{cases} X(t) = \varphi(t) + \int_0^t b(t,s,X(s))ds + \int_0^t \sigma(t,s,X(s))dW(s), \\ Y(t) = \psi(t) + \int_t^T g(t,s,X(s),Y(s),Z(s,t))ds - \int_t^T Z(t,s)dW(s), \end{cases} \tag{1}$$

*This work is supported by National Natural Science Foundation of China Grant 10771122, Natural Science Foundation of Shandong Province of China Grant Y2006A08 and National Basic Research Program of China (973 Program, No. 2007CB814900).

†E-mail: xiaotian2008001@gmail.com, yfshi@sdu.edu.cn

which for instance generalize the optimal control problems for stochastic Volterra integral equations in [26].

The notion of M-solution for backward stochastic Volterra integral equation (BSVIE in short) of the form

$$Y(t) = \psi(t) + \int_t^T g(t, s, Y(s), Z(s, t), Z(t, s))ds - \int_t^T Z(t, s)dW(s) \tag{2}$$

with $t \in [0, T]$, was introduced by Yong in [26], which plays an important role in optimal control problem for stochastic Volterra integral equations (SVIEs for short). We refer the author to Lin [7], Yong ([24], [25]), Wang and Zhang ([21]) for a study of the wellposedness of BSVIEs in finite space, while Anh and Yong [1], Ren [18] in infinite space counterpart.

One main feature of equation (2) lies in the dependence of the generator g on $Z(s, t)$, and hence it is quite different from, more precisely, a natural generalization of the one in [7] and [21]. Of course, the appearance of such term in g is not just means an extension from mathematical point of view, but also be of great importance in applications, (see Proposition 3.5 in [25] and Theorem 5.1 in [26].) It is interesting to realize that, so far as we know, it is just the term $Z(s, t)$ rather than $Z(t, s)$ in the generator g that plays a key role in both optimal control problem in [26] and dynamic risk measure in [25].

Optimal control of forward stochastic differential systems is a classical problem. When we consider the Pontryagin maximum principle for optimal controls of stochastic differential equations, the adjoint equation for variational state equation actually is a linear backward stochastic differential equation (BSDE for short). The wellposedness for nonlinear BSDEs was firstly studied by Pardoux and Peng [14]. Readers interested in an in-depth analysis of BSDEs can see the books of Ma and Yong [8], Yong and Zhou [28] and the survey paper of EI Karoui, Peng and Quenez [4]. As to the optimal control for stochastic differential equation, we refer the reader to, for example, Peng [15] for the general case of control domain being non-convex, and Yong and Zhou [28] for systematical analysis. On the other hand, optimal control for deterministic Volterra integral equation, particularly linear quadratic problem, was firstly studied by Vinokurov [19]. From then on some other extensions were developed, see, for example, [2], [3], [17], [27] and the references cited therein. As to the stochastic version, Yong ([24] and [26]) presented a maximum principle for SVIEs by means of BSVIEs, while the control is assumed to be convex. We also would like to mention the work of Øksendal and Zhang [13] in partial information setting without the help of BSVIEs. Along this, we will investigate the FBSVIEs case in this paper. To the best of our knowledge, so far little is known about maximum principle for FBSVIEs, and one aim of this manuscript is to close the gap.

The scheme is designed around the three steps for FBSDEs in Peng [16], namely listing out the variational equation, obtaining the variational inequality and utilizing some key mathematical tools to finish the procedure. As we know, within the context of stochastic differential systems, Itô formula has received most attention largely due to its ad hoc role in many complicated calculations and proofs. For example one usually

makes use of Itô formula in obtaining the convergence property for $\widetilde{X}_\rho$ and $\widetilde{Y}_\rho$ (defined blew) for differential systems, see Lemma 4.1 in [16]. In fact, one key tool in deriving maximum principle in Peng [16] is just Itô formula too. Unfortunately, this efficient tool is failure in the Volterra integral systems and some related well properties are absent in this case.

In this paper, new approaches are proposed to handle with the difficulties encountered in above procedure. On the one hand, we will make use of the dual principle, established by Yong in [26], for linear stochastic integral equation and its adjoint equation. Consequently we have to tackle four equations which perhaps means more mathematical expressions and notations involved after introducing another two more adjoint equations for FBSVIEs. As a result, it is our hope to choose appropriate form of adjoint equations so as to make the procedure as brief as possible. Fortunately, such adjoint equations really exist, see (10) and (22). On the other hand, we introduce a new equivalent norm for elements in $\mathcal{H}^2[0,T]$, see (6), and use some common calculations and tricks employed in the conventional BSDEs case, thereby obtain some convergence results, which play a chief role in deducing the variational inequality. Notice that Itô formula does not appear in the above two aspects.

Motivated by the new norm aforementioned, in the following we will provide a new method for the unique solvability of M-solution, which seems more convenient than the one in [26]. By the four steps in Theorem 3.7 in [26] we can see the process of constructing the M-solution clearly. From mathematical view, however, the whole proof is too complicated and uneasy to understand, which prompts us to seek an alternative one. We will carry out this course in detail in Section 3.

A class of continuous time dynamic convex and coherent risk measures, perhaps allowing time-inconsistent preference unlike the conventional case, were introduced by Yong in [25] via BSVIEs of the form

$$Y(t) = -\psi(t) + \int_t^T g(t,s,Y(s),Z(s,t))ds - \int_t^T Z(t,s)dW(s), \tag{3}$$

with $t \in [0,T]$. In the classical case, the terminal condition is usually a bounded random variable, representing the financial position at time T. However, in the situation under our consideration, we prefer to choose a process ψ, representing the total wealth of certain portfolio process at time t which might be a combination of certain contingent claims, positions of stocks, mutual funds and bonds. Usually the ψ could be $\mathcal{B}[0,T] \otimes \mathcal{F}_T$-measurable, see the example on p. 13 in [25]. If we define a map ϱ from $L^2_{\mathcal{F}_T}[0,T]$ to $L^2_{\mathbb{F}}[0,T]$ by $\varrho(t;\psi(\cdot)) = Y(t)$, with Y being the M-solution of BSVIE (3), given certain assumptions on g, it is shown in [25] that ϱ could be a dynamic convex/coherent risk measure. The question is how to look for a appropriate portfolio that minimizes the risk of the wealth process ψ by means of the representation above in finance, i.e., to seek an optimal solution for the so-called risk minimization problem, see Mataramvura and Øksendal [9], Øksendal and Sulem [12] for more information on the above financial problem. We conclude this paper by giving a study of this problem with the help of the maximum principle. In some cases, the closed form of optimal portfolio is derived.

The remainder of paper is organized as follows. In Section 2, we give some preliminary results and notations which are needed in the following sections. A new method for the solvability of M-solution is presented in Section 3. We give the stochastic maximum principle for FBSVIEs (1) as well as a backward linear quadratic problem in Section 4. At last, we investigate a risk minimization problem by means of maximum principle in the previous. Some explicit form solutions are derived.

2 Preliminaries

In this section, we will make some preliminaries. Let us specify some notation in this paper. For any $R, S \in [0,T]$, we denote $\Delta^c[0,T] = \{(t,s) \in [0,T]^2; t \leq s\}$, $\Delta[0,T] = \{(t,s) \in [0,T]^2; t > s\}$. In what follows some spaces will be frequently used. Let $L^p_{\mathcal{F}_T}[0,T]$ be the set of the $\mathcal{B}([0,T]) \otimes \mathcal{F}_T$ processes $X : [0,T] \times \Omega \to R^m$ such that $E\int_0^T |X(t)|^p dt < \infty$. $L^p_{\mathcal{F}}[0,T]$ is the set of all adapted processes $X : [0,T] \times \Omega \to R^m$ such that $E\int_0^T |X(s)|^p ds < \infty$. $L^p(0,T; L^2_{\mathcal{F}}[0,T])$ is the set of all processes $Z : [0,T]^2 \times \Omega \to R^{m\times d}$ such that for almost all $t \in [0,T]$, $Z(t,\cdot)$ is $\mathcal{F}$-progressively measurable satisfying $E\int_0^T \left(\int_0^T |Z(t,s)|^2 ds\right)^{\frac{p}{2}} dt < \infty$. For notational clarity, we denote $\mathcal{H}^p[0,T] = L^p_{\mathcal{F}}[0,T] \times L^p(0,T; L^2_{\mathcal{F}}[0,T])$. Next we shall cite the definition of M-solution introduced in [26].

Definition 2.1 *A pair of $(Y(\cdot), Z(\cdot,\cdot)) \in \mathcal{H}^p[0,T]$ is called an adapted M-solution of BSVIE (2) on $[0,T]$ if (2) holds in the usual Itô's sense for almost all $t \in [0,T]$ and, in addition, $Y(t) = EY(t) + \int_0^t Z(t,s)dW(s)$ with $t \in [0,T]$.*

The next two definitions are introduced by Yong in [25].

Definition 2.2 *A mapping $\rho : L^2_{\mathcal{F}_T}[0,T] \to L^2_{\mathbb{F}}[0,T]$ is called a dynamic risk measure if the following hold:*

1) (Past independence) For any $\Psi(\cdot)$, $\overline{\Psi}(\cdot) \in L^2_{\mathcal{F}_T}[0,T]$, if $\Psi(s) = \overline{\Psi}(s)$, a.s. $\omega \in \Omega$, $s \in [t,T]$, for some $t \in [0,T)$, then $\rho(t;\Psi(\cdot)) = \rho(t;\overline{\Psi}(\cdot))$, a.s. $\omega \in \Omega$.

2) (Monotonicity) For any $\Psi(\cdot)$, $\overline{\Psi}(\cdot) \in L^2_{\mathcal{F}_T}[0,T]$, if $\Psi(s) \leq \overline{\Psi}(s)$, a.s. $\omega \in \Omega$, $s \in [t,T]$, for some $t \in [0,T)$, then $\rho(s;\Psi(\cdot)) \geq \rho(s;\overline{\Psi}(\cdot))$, a.s. $\omega \in \Omega$, $s \in [t,T]$.

Definition 2.3 *A dynamic risk measure $\rho : L^2_{\mathcal{F}_T}[0,T] \to L^2_{\mathbb{F}}[0,T]$ is called a coherent risk measure if the following hold:*

1) There exists a deterministic integrable function $r(\cdot)$ such that for any $\Psi(\cdot) \in L^2_{\mathcal{F}_T}[0,T]$,

$$\rho(t;\Psi(\cdot) + c) = \rho(t;\Psi(\cdot)) - ce^{\int_t^T r(s)ds}, \ \text{a.s. } \omega \in \Omega, \ t \in [0,T].$$

2) For $\Psi(\cdot) \in L^2_{\mathcal{F}_T}[0,T]$ and $\lambda > 0$, $\rho(t;\lambda\Psi(\cdot)) = \lambda\rho(t;\Psi(\cdot))$ a.s. $\omega \in \Omega$, $t \in [0,T]$.

3) For any $\Psi(\cdot)$, $\overline{\Psi}(\cdot) \in L^2_{\mathcal{F}_T}[0,T]$,

$$\rho(t;\Psi(\cdot) + \overline{\Psi}(\cdot)) \leq \rho(t;\Psi(\cdot)) + \rho(t;\overline{\Psi}(\cdot)), \ \text{a.s. } \omega \in \Omega, \ t \in [0,T].$$

Some necessary specifications on the generator g for BSVIE (2) are given by:

(H1) Let $g : \Delta^c \times R^m \times R^{m\times d} \times R^{m\times d} \times \Omega \to R^m$ be $\mathcal{B}(\Delta^c \times R^m \times R^{m\times d} \times R^{m\times d}) \otimes \mathcal{F}_T$-measurable such that $s \to g(t,s,y,z,\zeta)$ is $\mathcal{F}$-progressively measurable for all $(t,y,z,\zeta) \in [0,T] \times R^m \times R^{m\times d} \times R^{m\times d}$, and $\forall y,\ \overline{y} \in R^m,\ z,\ \overline{z},\ \zeta,\ \overline{\zeta} \in R^{m\times d}$,

$$\begin{aligned} & |g(t,s,y,z,\zeta) - g(t,s,\overline{y},\overline{z},\overline{\zeta})| \\ \leq\ & L_1(t,s)|y-\overline{y}| + L_2(t,s)|z-\overline{z}| + L_3(t,s)|\zeta - \overline{\zeta}|, \end{aligned}$$

where $(t,s) \in \Delta^c$, $L_i(t,s)$ $(i = 1,2,3)$ is deterministic non-negative functions. Furthermore $E\int_0^T \left(\int_t^T |g_0(t,s)|ds\right)^p dt < \infty$, where $g_0(t,s) = g(t,s,0,0,0)$.

3 A new method for unique solvability of M-solution

In this section, a new scheme is proposed and analyzed to simplify the unique solvability of M-solution in Yong [26]. The proof in [26] gives us a detailed procedure to comprehend how to construct M-solutions, however, from a mathematical point of view, it is rather tedious and sophisticated, and it should be of interest to develop a new brief approach for it.

Inspired by the following equivalent norm for the elements of $\mathcal{H}^2[0,T]$ in [22],

$$\|(y(\cdot), z(\cdot,\cdot))\|_{\mathcal{H}^2[0,T]} = \left[E\int_0^T e^{\beta t}|y(t)|^2 dt + E\int_0^T \int_0^T e^{\beta s}|z(t,s)|^2 ds dt\right]^{\frac{1}{2}},$$

with β being a positive constant, we can propose a new one, see (6), and thus achieve the goal of giving a convenient and brief proof. In addition, compared with the proof in [22], it seems that the proof here is still simpler. Furthermore, we can also handle with the general case for $p \in (1,2]$ with this approach.

Before doing this, some preparations are required. Consider the following simple BSVIE,

$$Y(t) = \psi(t) + \int_t^T h(t,s,Z(t,s))ds - \int_t^T Z(t,s)dW(s). \tag{4}$$

(H2) h has the similar assumptions with g in (H1). Furthermore, $L_2(t,s)$ satisfies the condition, $\sup\limits_{t\in[0,T]} \int_t^T L_2(t,s)^{2+\epsilon} ds < \infty$, with some constant $\varepsilon > 0$.

The proof of the next proposition can be found in [26].

Proposition 3.1 *Let (H2) hold, then for any $\psi(\cdot) \in L^p_{\mathcal{F}_T}[0,T]$, (4) admits a unique adapted M-solution $(Y(\cdot), Z(\cdot,\cdot)) \in \mathcal{H}^p[0,T]$. If $\overline{h}$ also satisfies (H2), $\overline{\psi}(\cdot) \in L^p_{\mathcal{F}_T}[0,T]$, and $(\overline{Y}(\cdot), \overline{Z}(\cdot,\cdot)) \in \mathcal{H}^p[0,T]$ is the unique adapted M-solution of BSVIE (4) with (h,ψ) replaced by $(\overline{h}, \overline{\psi})$, then $\forall t \in [0,T]$,*

$$\begin{aligned} & E\left\{|Y(t) - \overline{Y}(t)|^p + \left(\int_t^T |Z(t,s) - \overline{Z}(t,s)|^2 ds\right)^{\frac{p}{2}}\right\} \\ \leq\ & CE\left[|\Psi(t) - \overline{\Psi}(t)|^p + \left(\int_t^T |h(t,s,Z(t,s)) - \overline{h}(t,s,Z(t,s))|ds\right)^p\right]. \end{aligned} \tag{5}$$

Hereafter C is a generic positive constant which may be different from line to line.

We move on to give the main result of this section.

Theorem 3.1 *Let (H1) hold, assume that*

$$\sup_{t\in[0,T]}\int_t^T L_1^q(t,s)ds<\infty,\quad \sup_{t\in[0,T]}\int_t^T L_2^{2+\epsilon}(t,s)ds<\infty,\quad \sup_{t\in[0,T]}\int_t^T L_3^{q'}(t,s)ds<\infty,$$

where $\frac{1}{p}+\frac{1}{q}=1$, $p\in(1,2]$, $\frac{1}{p'}+\frac{1}{q'}=1$, $1<p'<p$. Then for any $\psi(\cdot)\in L^p_{\mathcal{F}_T}[0,T]$, BSVIE (1) admits a unique adapted M-solution in $\mathcal{H}^p[0,T]$.

Proof. First let $\mathcal{M}^p[0,T]$ be the space of all $(y(\cdot),z(\cdot,\cdot))\in\mathcal{H}^p[0,T]$ such that $y(t)=Ey(t)+\int_0^t z(t,s)dW(s)$ with $t\in[0,T]$. It is a matter of direct calculation to show that $\mathcal{M}^p[0,T]$ is a closed subspace of $\mathcal{H}^p[0,T]$ via the following two martingale moment inequalities in [5],

$$E\int_0^T\left|\int_0^t z(t,s)dW(s)\right|^p dt\leq C_pE\int_0^T\left(\int_0^t|z(t,s)|^2ds\right)^{\frac{p}{2}}dt,\ \text{if } p>0,$$

and

$$E\int_0^T\left(\int_0^t|z(t,s)|^2ds\right)^{\frac{p}{2}}dt\leq C_pE\int_0^T\left|\int_0^t z(t,s)dW(s)\right|^p dt,\ \text{if } p>1,$$

where C_p is a constant depending on p. A new equivalent norm for the element in $\mathcal{M}^p[0,T]$ of the form

$$\|(y(\cdot),z(\cdot,\cdot))\|_{\mathcal{M}^p[0,T]}=\left[E\int_0^T e^{\beta t}|y(t)|^pdt+E\int_0^T e^{\beta t}\left(\int_0^T|z(t,s)|^2ds\right)^{\frac{p}{2}}dt\right]^{\frac{1}{p}},\quad (6)$$

will be in force in the following part. Consider,

$$Y(t)=\psi(t)+\int_t^T g(t,s,y(s),Z(t,s),z(s,t))ds-\int_t^T Z(t,s)dW(s),\qquad (7)$$

with $t\in[0,T]$, $\psi(\cdot)\in L^p_{\mathcal{F}_T}[0,T]$ and $(y(\cdot),z(\cdot,\cdot))\in\mathcal{M}^p[0,T]$. Following Proposition 3.1 we get that (7) admits a unique adapted M-solution $(Y(\cdot),Z(\cdot,\cdot))$, and then define a map $\Theta:\mathcal{M}^p[0,T]\to\mathcal{M}^p[0,T]$ by

$$\Theta(y(\cdot),z(\cdot,\cdot))=(Y(\cdot),Z(\cdot,\cdot)),\quad \forall(y(\cdot),z(\cdot,\cdot))\in\mathcal{M}^p[0,T].$$

Let $(\overline{y}(\cdot),\overline{z}(\cdot,\cdot))\in\mathcal{M}^p[0,T]$ and $\Theta(\overline{y}(\cdot),\overline{z}(\cdot,\cdot))=(\overline{Y}(\cdot),\overline{Z}(\cdot,\cdot))$, thus it follows from

inequality (5) that,

$$
\begin{aligned}
& E\int_0^T e^{\beta t}|Y(t)-\overline{Y}(t)|^p dt + E\int_0^T e^{\beta t}\left(\int_t^T |Z(t,s)-\overline{Z}(t,s)|^2 ds\right)^{\frac{p}{2}} dt \\
\leq\ & CE\int_0^T e^{\beta t}\left\{\int_t^T |g(t,s,y(s),Z(t,s),z(s,t))-g(t,s,\overline{y}(s),Z(t,s),\overline{z}(s,t))|ds\right\}^p dt \\
\leq\ & CE\int_0^T e^{\beta t}\left\{\int_t^T L_1(t,s)|y(s)-\overline{y}(s)|ds\right\}^p dt \\
& +CE\int_0^T e^{\beta t}\left\{\int_t^T L_3(t,s)|z(s,t)-\overline{z}(s,t)|ds\right\}^p dt \\
\leq\ & CE\int_0^T e^{\beta t}\left(\sup_{t\in[0,T]}\int_t^T L_1^q(t,s)ds\right)^{\frac{p}{q}}\int_t^T |y(s)-\overline{y}(s)|^p dsdt \\
& +CE\int_0^T e^{\beta t}\left(\sup_{t\in[0,T]}\int_t^T L_3^{q'}(t,s)ds\right)^{\frac{p}{q'}}\left(\int_t^T |z(s,t)-\overline{z}(s,t)|^{p'}ds\right)^{\frac{p}{p'}} dt \\
\leq\ & CE\int_0^T |y(s)-\overline{y}(s)|^p ds\int_0^s e^{\beta t}dt + C\left[\frac{1}{\beta}\right]^{\frac{p-p'}{p'}} E\int_0^T ds\int_t^T e^{\beta s}|z(s,t)-\overline{z}(s,t)|^p dt \\
\leq\ & \frac{C}{\beta}E\int_0^T e^{\beta s}|y(s)-\overline{y}(s)|^p ds + C\left[\frac{1}{\beta}\right]^{\frac{p-p'}{p'}} E\int_0^T e^{\beta t}dt\int_0^t |z(t,s)-\overline{z}(t,s)|^p ds \\
\leq\ & \frac{C}{\beta}E\int_0^T e^{\beta s}|y(s)-\overline{y}(s)|^p ds,
\end{aligned}
$$

where $1<p'<p$, $\frac{1}{p'}+\frac{1}{q'}=1$. Notice that here we use the following two relations, that are, for any $p\in(1,2]$, $1<p'<p$, $r>0$,

$$
\begin{aligned}
& \left[\int_t^T |z(s,t)-\overline{z}(s,t)|^{p'}ds\right]^{\frac{p}{p'}} \\
\leq\ & \left[\int_t^T e^{-rs\frac{p}{p-p'}}ds\right]^{\frac{p-p'}{p'}}\int_t^T e^{rs\frac{p}{p'}}|z(s,t)-\overline{z}(s,t)|^p ds \\
\leq\ & \left[\frac{1}{r}\right]^{\frac{p-p'}{p'}}\left[\frac{p-p'}{p}\right]^{\frac{p-p'}{p'}} e^{-rt\frac{p}{p'}}\int_t^T e^{rs\frac{p}{p'}}|z(s,t)-\overline{z}(s,t)|^p ds. \qquad (8)
\end{aligned}
$$

and $E\int_0^t |z(t,s)-\overline{z}(t,s)|^p ds \leq CE|y(t)-\overline{y}(t)|^p$ which is a direct consequence of martingale moment inequality and Hölder inequality. Then we can choose a β, so that the map Θ is a contraction, and the result holds. □

4 A maximum principle for FBSVIE

In this section, we give a stochastic maximum principle for forward-backward stochastic Volterra integral equations by assuming the control domain being convex and $p=2$,

thereby generalizing for instance the case in [26]. As compared with the differential case, it is by no means clear that the method there can be extended to such setting. As we have claimed in the previous, there are some technical obstacles for us to overcome due to the absence of Itô formula here, in other words, we should adopt some other effective mathematical skills to circumvent the difficulties caused by it. Without loss of generality, we assume that $m = d = 1$.

4.1 Setting the problem

We denote by U a nonempty convex subset of R, and set

$$\mathcal{U} = \{v(\cdot) \in L^2_{\mathcal{F}}[0,T]; v(t) \in U, a.s. \quad t \in [0,T], a.e.\}.$$

An element of $\mathcal{U}$ is called an admissible control. For any admissible control $v(\cdot) \in \mathcal{U}$, let us consider the following forward-backward stochastic Volterra integral equation, i.e.,

$$\begin{cases} X(t) = \varphi(t) + \int_0^t b(t,s,X(s),v(s))ds + \int_0^t \sigma(t,s,X(s),v(s))dW(s), \\ Y(t) = \psi(t) + \int_t^T g(t,s,X(s),Y(s),Z(s,t),v(s))ds - \int_t^T Z(t,s)dW(s), \end{cases} \tag{9}$$

associated with the cost functional by

$$J(v(\cdot)) = E\left[\int_0^T l(s,X(s),Y(s),v(s))ds + h(X(T)) + \gamma(Y(0))\right],$$

where $\varphi(\cdot) \in L^2_{\mathcal{F}}[0,T]$ and $\psi(\cdot) \in L^2_{\mathcal{F}_T}[0,T]$. Basic assumptions imposed on b, σ, g, l, h, γ are stated as

(H3)

$$\begin{aligned} b(t,s,x,v) &: \Delta \times R \times U \times \Omega \to R, \\ \sigma(t,s,x,v) &: \Delta \times R \times U \times \Omega \to R, \\ g(t,s,x,y,z,v) &: \Delta^c \times R \times R \times R \times U \times \Omega \to R, \\ l(s,x,y,v) &: [0,T] \times R \times R \times U \times \Omega \to R, \end{aligned}$$

$h(x) : \Omega \times R \to R$, $\gamma(x) : \Omega \times R \to R$. b, σ, g, l, h, and γ are continuously differentiable with respect to the variables. The derivatives of b, σ, g are bounded, the derivatives of l are bounded by $C(1 + |x| + |y| + |v|)$ and the derivatives of h, γ with respect to x are bounded by $C(1 + |x|)$. Furthermore, we assume that $g_i(t,s,x,y,z,v)$ is $\mathcal{B}(\Delta^c \times R \times R \times R \times U) \otimes \mathcal{F}_T$-measurable such that $t \mapsto g_i(t,s,x,y,z,v)$ is $\mathcal{F}$-progressively measurable for all $(s,x,y,z,v) \in [0,T] \times R \times R \times R \times U$, $(i = y, z)$.

Given (H3) and $v \in \mathcal{U}$, we observe that there exists a unique adapted M-solution $(X(\cdot), Y(\cdot), Z(\cdot,\cdot)) \in L^2_{\mathcal{F}}[0,T] \times L^2_{\mathcal{F}}[0,T] \times L^2(0,T; L^2_{\mathcal{F}}[0,T])$ for above FBSVIE (9) by what we mean that $X(\cdot)$ satisfies the forward equation in (9) in the usual sense

and $(Y(\cdot), Z(\cdot,\cdot))$ is the adapted M-solution of the backward form of (9). Both of the equations in (9) are called the state equations. The optimal control problem is to minimize the cost function $J(v(\cdot))$ over admissible controls. An admissible control $v(\cdot)$ is called an optimal control if it attains the minimum.

Remark 4.1 *A special case of the above optimal control problem was considered in [26] where*

$$J(u(\cdot)) = E\left[\int_0^T l(s, X(s), v(s))ds + h(X(T))\right],$$

and the coefficients are assumed to be independent of ω.

4.2 Variational equations and one convergence result

Let $u(\cdot)$ be an optimal control and let $(X(\cdot), Y(\cdot), Z(\cdot,\cdot))$ be the corresponding M-solution of (9). Let $v(\cdot)$ be such that $u(\cdot)+v(\cdot) \in \mathcal{U}$. Since $\mathcal{U}$ is convex, then for any $0 \le \rho \le 1$, $u_\rho = u(\cdot) + \rho v(\cdot) \in \mathcal{U}$. Let's us consider,

$$\begin{cases} \xi(t) = \varphi_1(t) + \displaystyle\int_0^t b_x(t,s,X^u(s),u(s))\xi(s)ds + \int_0^t \sigma_x(t,s,X^u(s),u(s))\xi(s)dW(s), \\ \eta(t) = \psi_1(t) + \displaystyle\int_t^T g_y(t,s,X^u(s),Y^u(s),Z^u(s,t),u(s))\eta(s)ds \\ \qquad + \displaystyle\int_t^T g_z(t,s,X^u(s),Y^u(s),Z^u(s,t),u(s))\zeta(s,t)ds - \int_t^T \zeta(t,s)dW(s), \end{cases} \tag{10}$$

where

$$\begin{cases} \varphi_1(t) = \displaystyle\int_0^t b_v(t,s,X^u(s),u(s))v(s)ds + \int_0^t \sigma_v(t,s,X^u(s),u(s))v(s)dW(s), \\ \psi_1(t) = \displaystyle\int_t^T g_x(t,s,X^u(s),Y^u(s),Z^u(s,t),u(s))\xi(s)ds \\ \qquad + \displaystyle\int_t^T g_v(t,s,X^u(s),Y^u(s),Z^u(s,t),u(s))v(s)ds. \end{cases} \tag{11}$$

The two equations in (10) are called variational equations. Obviously under assumption (H3) we can find a unique M-solution $(\xi(\cdot), \eta(\cdot), \zeta(\cdot,\cdot)) \in L^2_{\mathcal{F}}[0,T] \times L^2_{\mathcal{F}}[0,T] \times L^2(0,T; L^2_{\mathcal{F}}[0,T])$, which is the unique adapted M-solution of FBSVIE (10). We denote by $(X_\rho(\cdot), Y_\rho(\cdot), Z_\rho(\cdot,\cdot))$ the M-solutions of (9) corresponding to u_ρ. We now proceed to prove the relations

$$\begin{cases} E\displaystyle\int_0^T |X_\rho(t) - X^u(t)|^2 dt \to 0; \quad \rho \to 0, \\ E\displaystyle\int_0^T |Y_\rho(t) - Y^u(t)|^2 dt \to 0; \quad \rho \to 0, \\ E\displaystyle\int_0^T \int_0^t |Z_\rho(t,s) - Z^u(t,s)|^2 dsdt \to 0; \quad \rho \to 0. \end{cases} \tag{12}$$

In fact, it follows from the denotation of X_ρ, together with the forward equation in (9) that

$$\begin{aligned}
& E\int_0^T e^{-\gamma t}|X_\rho(t)-X^u(t)|^2dt \\
\leq\ & CE\int_0^T e^{-\gamma t}dt\int_0^t |X_\rho(s)-X^u(s)|^2ds + CE\int_0^T e^{-\gamma t}dt\int_0^t |u_\rho(s)-u(s)|^2ds \\
\leq\ & CE\int_0^T |X_\rho(s)-X^u(s)|^2ds\int_s^T e^{-\gamma t}dt + CE\int_0^T |u_\rho(s)-u(s)|^2ds\int_s^T e^{-\gamma t}dt \\
\leq\ & \frac{C}{\gamma}E\int_0^T e^{-\gamma s}|X_\rho(s)-X^u(s)|^2ds + \frac{C}{\gamma}E\int_0^T e^{-\gamma s}|u_\rho(s)-u(s)|^2ds, \qquad (13)
\end{aligned}$$

where γ is a positive constant depending on the upper bound of all the derivatives. By choosing a γ such that $\frac{C}{\gamma}=\frac{1}{2}$, it leads to

$$E\int_0^T |X_\rho(t)-X^u(t)|^2dt \leq e^{\gamma T}E\int_0^T e^{-\gamma t}|X_\rho(t)-X^u(t)|^2dt \to 0; \quad \rho\to 0.$$

Following the conclusion of Theorem 3.7 in [26], we observe that

$$\begin{aligned}
& E\int_0^T |Y_\rho(t)-Y^u(t)|^2dt + E\int_0^T\int_0^T |Z_\rho(t,s)-Z^u(t,s)|^2dsdt \\
\leq\ & CE\int_0^T \left(\int_t^T |g'(t,s,Y_\rho(s),Z_\rho(s,t)) - g''(t,s,Y_\rho(s),Z_\rho(s,t))|ds\right)^2 dt \\
\leq\ & CE\int_0^T |X_\rho(s)-X^u(s)|^2ds + CE\int_0^T |u_\rho(s)-u(s)|^2ds \to 0, \quad \rho\to 0, \quad (14)
\end{aligned}$$

where $g'(t,s,y,z)=g(t,s,X_\rho(s),y,z,u_\rho(s))$, $g''(t,s,y,z)=g(t,s,X(s),y,z,u(s))$, C is a constant depending on the upper bound of all the derivatives. Thus we can get (12). For $t,s\in[0,T]$, set

$$\left\{\begin{array}{l}
\widetilde{X}_\rho(t)=\rho^{-1}(X_\rho(t)-X^u(t))-\xi(t), \\
\widetilde{Y}_\rho(t)=\rho^{-1}(Y_\rho(t)-Y^u(t))-\eta(t), \\
\widetilde{Z}_\rho(s,t)=\rho^{-1}(Z_\rho(s,t)-Z^u(s,t))-\zeta(s,t).
\end{array}\right. \qquad (15)$$

Using the similar method as (13), recalling the denotation of X_ρ, we can deduce that

$$E\int_0^T e^{-\alpha t}|\widetilde{X}_\rho(t)|^2dt \leq \frac{C}{\alpha}E\int_0^T e^{-\alpha t}|\widetilde{X}_\rho(t)|^2dt + \varepsilon_\rho,$$

where C is a constant depending on the upper bound of the derivatives, and $\varepsilon_\rho\to 0$, $\rho\to 0$. Then we can choose α such that $\frac{C}{\alpha}=\frac{1}{2}$, and

$$E\int_0^T |\widetilde{X}_\rho(t)|^2dt \leq e^{\alpha T}E\int_0^T e^{-\alpha t}|\widetilde{X}_\rho(t)|^2dt \leq 2e^{\alpha T}\varepsilon_\rho \to 0; \quad \rho\to 0.$$

As to the term $\widetilde{Y}_\rho$, we arrive at

$$
\begin{aligned}
& E\int_0^T e^{\beta t}|\widetilde{Y}_\rho(t)|^2dt + E\int_0^T e^{\beta t}\int_t^T |\widetilde{Z}_\rho(t,s)|^2dsdt \\
\leq\ & CE\int_0^T e^{\beta t}\int_t^T |\widetilde{X}_\rho(s)|^2dsdt + CE\int_0^T e^{\beta t}\int_t^T |\widetilde{Y}_\rho(s)|^2dsdt \\
& +\frac{C}{\beta}E\int_0^T\int_t^T e^{\beta s}|\widetilde{Z}_\rho(s,t)|^2dsdt + Ce^{\beta T}\varepsilon_\rho' \\
\leq\ & \frac{C}{\beta}E\int_0^T e^{\beta s}|\widetilde{X}_\rho(s)|^2ds + \frac{C}{\beta}E\int_0^T e^{\beta s}|\widetilde{Y}_\rho(s)|^2ds + Ce^{\beta T}\varepsilon_\rho',
\end{aligned}
$$

where C is an constant depending on the upper bound of all the derivative, and $\varepsilon_\rho' \to 0$, $\rho \to 0$. Then we can choose a β so that $\frac{C}{\beta} < 1$, and

$$
\begin{aligned}
& E\int_0^T e^{\beta t}|\widetilde{Y}_\rho(t)|^2dt + E\int_0^T e^{\beta t}\int_t^T |\widetilde{Z}_\rho(t,s)|^2dsdt \\
\leq\ & CE\int_0^T e^{\beta s}|\widetilde{X}_\rho(s)|^2ds + Ce^{\beta T}\varepsilon_\rho'.
\end{aligned}
$$

From above

$$
CE\int_0^T e^{\beta s}|\widetilde{X}_\rho(s)|^2ds \leq Ce^{\beta T}E\int_0^T |\widetilde{X}_\rho(t)|^2dt \to 0; \quad \rho \to 0,
$$

thus

$$
E\int_0^T |\widetilde{Y}_\rho(t)|^2dt \leq E\int_0^T e^{\beta t}|\widetilde{Y}_\rho(t)|^2dt \to 0; \quad \rho \to 0.
$$

To sum up the argument above, we obtain:

Lemma 4.1 *Let (H3) hold, then*

$$
\lim_{\rho\to 0} E\int_0^T |\widetilde{X}_\rho(s)|^2ds = 0, \quad \lim_{\rho\to 0} E\int_0^T |\widetilde{Y}_\rho(s)|^2ds = 0. \tag{16}
$$

4.3 A simple form of stochastic maximum principle

In what follows, we make the following conventions with $t, s \in [0,T]$, $v \in \mathcal{U}$,

$$
\begin{aligned}
& l_i^v(s) = l_i(s, X^v(s), Y^v(s), v(s)), i = x, y, v, \\
& h_j^v(s,t) = h_j(s,t,X^v(t), v(t)), j = x, v, h = b, \sigma, \\
& g_k^v(s,t) = g_k(s,t,X^v(t), Y^v(t), Z^v(t,s), v(t)), k = x, y, z, v,
\end{aligned}
$$

where (X^v, Y^v, Z^v) is the solution of (9) corespondent to v. In this subsection we assume that the cost function takes a simple form of $J(u(\cdot)) = E\int_0^T l(s, X(s), Y(s), u(s))ds$. Since u is an optimal control, then $\rho^{-1}[J(u+\rho v) - J(u)] \geq 0$, and we have the following variational inequality.

Lemma 4.2 *Let (H3) hold, then*

$$E\int_0^T l_x^u(s)\xi(s)ds + E\int_0^T l_y^u(s)\eta(s)ds + E\int_0^T l_v^u(s)v(s)ds \geq 0, \tag{17}$$

where (X, Y, Z) is the unique M-solution of FBSVIE (9) with u being an optimal control.

Proof. From the Lemma 4.1, we know

$$\begin{aligned}
& \rho^{-1}E\int_0^T [l(s, X_\rho(s), Y_\rho(s), u_\rho(s)) - l(s, X^u(s), Y^u(s), u(s))]ds \\
= \quad & E\int_0^T l_x(s, X^u(s) + \theta(X_\rho(s) - X^u(s)), Y_\rho(s), u_\rho(s))\frac{X_\rho(s) - X^u(s)}{\rho}ds \\
& + E\int_0^T l_y(s, X^u(s), Y^u(s) + \theta(Y_\rho(s) - Y^u(s)), u_\rho(s))\frac{Y_\rho(s) - Y^u(s)}{\rho}ds \\
& + E\int_0^T l_u(s, X^u(s), Y^u(s), u(s) + \theta(u_\rho(s) - u(s)))v(s)ds \\
\to \quad & E\int_0^T l_x(s, X^u(s), Y^u(s), u(s))\xi(s)ds + E\int_0^T l_y(s, X^u(s), Y^u(s), u(s))\eta(s)ds \\
& + E\int_0^T l_v(s, X^u(s), Y^u(s), u(s))v(s)ds.
\end{aligned}$$

Thus the conclusion follows. $\square$

For deriving the maximum principle, we introduce the following two adjoint equations:

$$\begin{cases}
P(t) = l_y^u(t) + \displaystyle\int_0^t g_y^u(s,t)P(s)ds + \int_0^t g_z^u(s,t)P(s)dW(s), \\
Q(t) = l_x^u(t) + \displaystyle\int_0^t g_x^u(s,t)P(s)ds + \int_t^T b_x^u(s,t)Q(s)ds \\
\qquad\qquad + \displaystyle\int_t^T \sigma_x^u(s,t)R(s,t)ds - \int_t^T R(t,s)dW(s).
\end{cases} \tag{18}$$

Obviously the above FBSVIE admits a unique M-solution $(P(\cdot), Q(\cdot), R(\cdot,\cdot))$ under assumption (H3). Note that g_y and g_z are non-anticipated processes under (H3). The later proposition is the so-called dual principle for linear stochastic Volterra integral equation, the proof of which can be found in [26].

Proposition 4.1 *Let $A_i : \triangle \times \Omega \to R$ $(i = 1, 2)$ be $\mathcal{B}(\triangle) \otimes \mathcal{F}_T$-measurable such that $s \to A(t, s)$ is $\mathcal{F}$-progressively measurable for all $t \in [0, T]$, furthermore, we assume that they are two bounded processes, $\varphi(\cdot) \in L^2_{\mathcal{F}}[0, T]$ and $\psi(\cdot) \in L^2_{\mathcal{F}_T}[0, T]$. Let $\xi(\cdot) \in L^2_{\mathcal{F}}[0, T]$ be the solution of FSVIE:*

$$\xi(t) = \varphi(t) + \int_0^t A_1(t,s)\xi(s)ds + \int_0^t A_2(t,s)\xi(s)dW(s), \quad t \in [0, T],$$

and $(Y(\cdot), Z(\cdot,\cdot))$ *be the adapted M-solution to the following BSVIE,* $\forall t \in [0,T]$,

$$Y(t) = \psi(t) + \int_t^T \{A_1(s,t)Y(s) + A_2(s,t)Z(s,t)\}ds - \int_t^T Z(t,s)dW(s).$$

Then the following relation holds:

$$E\int_0^T \xi(t)\psi(t)dt = E\int_0^T \varphi(t)Y(t)dt.$$

We now assert:

Theorem 4.1 *Let* $u(\cdot)$ *be an optimal control and* $(X^u(\cdot), Y^u(\cdot), Z^u(\cdot,\cdot))$ *be the corresponding M-solution of FBSVIE (9). Then we have,* $\forall v \in U$,

$$H(X^u(t), Y^u(t), Z^u(t,\cdot), u(t), P(t), Q(t), R(\cdot,t)) \cdot (v - u(t)) \geq 0, \quad a.e., a.s.$$

where

$$\begin{aligned}
& H(X^u(t), Y^u(t), Z^u(t,\cdot), u(t), P(t), Q(t), R(\cdot,t)) \\
= \; & l_v^u(t) + E^{\mathcal{F}_t}\int_t^T Q(s)b_v^u(s,t)ds + E^{\mathcal{F}_t}\int_t^T R(s,t)\sigma_v^u(s,t)ds + \int_0^t g_v^u(s,t)P(s)ds
\end{aligned}$$

Here (P,Q,R) *is the unique M-solution of FBSVIE (18).*

Proof. From the forward form in (10), the backward form in (18) and Proposition 4.1 above, we know that

$$\begin{aligned}
& E\int_0^T \xi(t)l_x^u(t)dt + E\int_0^T \xi(t)\int_0^t g_x^u(s,t)P(s)dsdt \\
= \; & E\int_0^T Q(t)dt\int_0^t b_v^u(t,s)v(s)ds + E\int_0^T Q(t)dt\int_0^t \sigma_v^u(t,s)v(s)dW(s) \\
= \; & E\int_0^T Q(t)dt\int_0^t b_v^u(t,s)v(s)ds + E\int_0^T\int_0^t R(t,s)\sigma_v^u(t,s)v(s)dsdt \\
= \; & E\int_0^T v(s)K(s)ds,
\end{aligned} \tag{19}$$

where

$$K(s) = \int_s^T \{Q(t)b_v^u(t,s) + R(t,s)\sigma_v^u(t,s)\}dt.$$

Similarly from the backward form in (10), the forward form in (18) and Proposition 4.1, we know

$$\begin{aligned}
E\int_0^T \eta(t)l_y^u(t)dt &= E\int_0^T P(t)dt\int_t^T g_x^u(t,s)\xi(s)ds + E\int_0^T P(t)dt\int_t^T g_v^u(t,s)v(s)ds \\
&= E\int_0^T \xi(t)dt\int_0^t g_x^u(s,t)P(s)ds + E\int_0^T v(s)ds\int_0^s g_v^u(t,s)P(t)dt.
\end{aligned} \tag{20}$$

It follows form (19) and (20),

$$\begin{aligned}
&E\int_0^T \xi(t)l_x^u(t)dt + E\int_0^T \eta(t)l_y^u(t)dt \\
= \quad & E\int_0^T v(s)ds\left[\int_s^T Q(t)b_v^u(t,s)dt + \int_s^T R(t,s)\sigma_v^u(t,s)dt + \int_0^s g_v^u(t,s)P(t)dt\right].
\end{aligned}$$

From the variational inequality (17) we have

$$\begin{aligned}
0 \quad \leq \quad & E\int_0^T l_x^u(t)\xi(t)dt + E\int_0^T l_y^u(t)\eta(t)dt + E\int_0^T l_v^u(t)v(t)dt \\
= \quad & E\int_0^T v(t)L(t)dt,
\end{aligned}$$

where

$$L(t) \quad = \quad l_v^u(t) + \int_t^T Q(s)b_v^u(s,t)ds + \int_t^T R(s,t)\sigma_v^u(s,t)ds + \int_0^t g_v^u(s,t)P(s)ds.$$

The proof is complete. $\square$

4.4 A general stochastic maximum principle

In this subsection we will give a new maximum principle, while the cost function is a more general form

$$J(v(\cdot)) = E\left[\int_0^T l(s,X(s),Y(s),v(s))ds + h(X(T)) + \gamma(Y(0))\right].$$

It can be easily checked that $E\int_0^T l_x^u(s)\xi(s)ds + E\int_0^T l_y^u(s)\eta(s)ds + E\int_0^T l_v^u(s)v(s)ds + Eh_x(X^u(T))\xi(T)+E\gamma_y(Y^u(0))\eta(0) \geq 0$. In fact, the definition of $\widetilde{X}_\rho$ implies $E|\widetilde{X}_\rho(T)|^2 \leq \delta_1(\rho) + CE\int_0^T |\widetilde{X}_\rho(s)|^2ds$, with $\delta_1(\rho) \to 0, \rho \to 0$, C is a constant depending on the upper bound of all the derivatives. Recalling the result in Lemma 4.1 we obtain $E|\widetilde{X}_\rho(T)|^2 \to 0, \rho \to 0$. Similarly by the form of $\widetilde{Y}_\rho$, it follows that

$$E|\widetilde{Y}_\rho(0)|^2 \leq CE\int_0^T |\widetilde{X}_\rho(s)|^2ds + CE\int_0^T |\widetilde{Y}_\rho(s)|^2ds + CE\int_0^T |\widetilde{Z}_\rho(s,0)|^2ds + \delta_2(\rho),$$

with $\delta_2(\rho) \to 0, \rho \to 0$. By the definition of M-solution in the previous, it does no matter what value of $\zeta(s,0)$ is as long as it is a deterministic function on $s \in [0,T]$. In particular, if $E\int_0^t |D_s\eta(t)|^2ds < \infty$ (here D is a malliavin operator, see [11] for more detailed accounts for malliavin calculus), then by Ocone-Clark formula (see [11]) and the definition of M-solution, we obtain

$$\eta(t) = E\eta(t) + \int_0^t \zeta(t,s)dW(s) = E\eta(t) + \int_0^t E^{\mathcal{F}_s}D_s\eta(t)dW(s),$$

thus we have $\zeta(t,s) = E^{\mathcal{F}_s} D_s \eta(t)$, then without loss of generality we can determine $\zeta(s,0)$ by $E\eta(s)$. Similarly $\widetilde{Z}_\rho(s,0) = E\widetilde{Y}_\rho(s)$, and this leads to $E|\widetilde{Y}_\rho(0)|^2 \to 0$, with $\rho \to 0$.

Summing up, we finally obtain $\rho^{-1}E(h(X_\rho(T)) - h(X^u(T))) \to Eh_x(X^u(T))\xi(T)$, and $\rho^{-1}E(\gamma(Y_\rho(0)) - \gamma(Y^u(0))) \to E\gamma_y(Y^u(0))\eta(0)$, with $E|\xi(T)|^2 < \infty, \quad E|\eta(0)|^2 < \infty$, which is easy to validate.

It follows from the martingale representation theorem that there exists a unique process $\pi(s) \in L^2_{\mathcal{F}}[0,T]$ so that $h_x(X^u(T)) = Eh_x(X^u(T)) + \int_0^T \pi(s)dW(s)$; then

$$
\begin{aligned}
& Eh_x(X^u(T))\xi(T) \\
= \;& Eh_x(X^u(T))\left[\int_0^T b_v^u(T,s)v(s)ds + \int_0^T \sigma_v^u(T,s)v(s)dW(s)\right] \\
& + Eh_x(X^u(T))\left[\int_0^T b_x^u(T,s)\xi(s)ds + \int_0^T \sigma_x^u(T,s)\xi(s)dW(s)\right] \\
= \;& E\int_0^T b_v^u(T,s)h_x(X^u(T))v(s)ds + E\int_0^T \pi(s)\sigma_v^u(T,s)v(s)ds \\
& + E\int_0^T b_x^u(T,s)\xi(s)h_x(X^u(T))ds + E\int_0^T \sigma_x^u(T,s)\xi(s)\pi(s)ds.
\end{aligned}
$$

On the other hand, using the fact that $E\gamma_y(Y^u(0))\int_0^T \zeta(0,s)dW(s) = 0$, one gets

$$
\begin{aligned}
E\gamma_y(Y^u(0))\eta(0) \;=\;& E\int_0^T g_x^u(0,s)\gamma_y(Y^u(0))\xi(s)ds + E\int_0^T g_v^u(0,s)\gamma_y(Y^u(0))v(s)ds \\
+\;& E\int_0^T g_y^u(0,s)\gamma_y(Y^u(0))\eta(s)ds + E\int_0^T g_z^u(0,s)\gamma_y(Y^u(0))E\eta(s)ds.
\end{aligned}
\tag{21}
$$

In this case, FBSVIE (18) is replaced by

$$
\left\{
\begin{aligned}
P(t) &= l_y^u(t) + g_y^u(0,t)\gamma_y(Y^u(0)) + \gamma_y(Y^u(0))Eg_z^u(0,t) \\
&\quad + \int_0^t g_y^u(s,t)P(s)ds + \int_0^t g_z^u(s,t)P(s)dW(s), \\
Q(t) &= l_x^u(t) + b_x^u(T,t)h_x(X^u(T)) + \sigma_x^u(T,t)\pi(t) + g_x^u(0,t)\gamma_y(Y^u(0)) + \int_0^t g_x^u(s,t)P(s)ds \\
&\quad + \int_t^T b_x^u(s,t)Q(s)ds + \int_t^T \sigma_x^u(s,t)R(s,t)ds - \int_t^T R(t,s)dW(s).
\end{aligned}
\right.
\tag{22}
$$

So by a similar proof as Theorem 4.1 we get a more general stochastic maximum principle.

Theorem 4.2 *Let $u(\cdot)$ be an optimal control and $(X(\cdot), Y(\cdot), Z(\cdot,\cdot))$ be the corresponding M-solution of FBSVIE (9). Then we have, $\forall v \in U$,*

$$
H(t, X^u(t), Y^u(t), Z^u(t,\cdot), u(t), P(t), Q(t), R(\cdot,t)) \cdot (v - u(t)) \geq 0, \quad a.e., a.s. \tag{23}
$$

where

$$
\begin{aligned}
& H(t, X(t), Y(t), Z(t, \cdot), u(t), P(t), Q(t), R(\cdot, t)) \\
= \;& l_v^u(t) + b_v^u(T, t) E^{\mathcal{F}_t} h_x(X^u(T)) + \sigma_v^u(T, t)\pi(t) + g_v^u(0, t) \\
& + \int_0^t g_v^u(s, t) P(s) ds + E^{\mathcal{F}_t} \int_t^T R(s, t)\sigma_v^u(s, t) ds + E^{\mathcal{F}_t} \int_t^T Q(s) b_v^u(s, t) ds
\end{aligned}
$$

where (P, Q, R) is the unique M-solution of FBSVIE (22).

If we define

$$
\mathcal{H}(t, X^u(t), Y^u(t), Z^u(t, \cdot), u(t), P(t), Q(t), R(\cdot, t), v) = -H \cdot v,
$$

then (23) can be rewritten as

$$
\begin{aligned}
& \mathcal{H}(t, X^u(t), Y^u(t), Z^u(t, \cdot), u(t), P(t), Q(t), R(\cdot, t), u(t)) \\
= \;& \max_{v \in U} \mathcal{H}(t, X^u(t), Y^u(t), Z^u(t, \cdot), u(t), P(t), Q(t), R(\cdot, t), v)
\end{aligned} \tag{24}
$$

We call $\mathcal{H}$ the Hamiltonian of the optimal control problem of FBSVIEs, and (24) the maximum principle condition.

We would like to conclude this section by giving a application, that is, a linear-quadratic (LQ for short) problem of BSVIEs. The linear BSVIE is of the form

$$
Y(t) = \psi(t) + \int_t^T [l_1(t, s)Y(s) + l_2(t, s)v(s) + l_3(t, s)Z(s, t)] ds - \int_t^T Z(t, s) dW(s), \tag{25}
$$

while the cost functional $J(\psi(\cdot), v(\cdot))$ associated with the terminal condition $\psi(\cdot)$ and control $v(\cdot)$ is given by

$$
J(\psi(\cdot), v(\cdot)) = \frac{1}{2} E \int_0^T [Q'(t)Y^2(t) + R'(t)v^2(t)] dt + \frac{1}{2} E G' Y^2(0).
$$

The linear-quadratic control problem is to minimize the cost function over admissible controls. Necessary assumptions will be in force in the following.

(H4) Let $l_i : \Delta^c \to R$, $(i = 1, 2, 3)$ be three continuous bounded processes such that $s \to l_i(t, s)$ is $\mathbb{F}$ adapted for $t \in [0, T]$. Q' and R are bounded and non-negative adapted processes, moreover, $R'(t) > \delta$, where δ is a positive constant, G' is a non-negative bounded random variable, $\psi(\cdot) \in L^2_{\mathcal{F}_T}[0, T]$. In addition, assume U is also closed.

Obviously (H4) is sufficient for the finiteness of the above linear-quadratic problem. Following the idea of Theorem 5.2 in Chapter 2 of [28], we are ready to present a existence theorem of the optimal control.

Lemma 4.3 *Let (H4) hold, then there exists a $u(\cdot) \in \mathcal{U}$ such that $J(\psi, u) = \inf_{v \in \mathcal{U}} J(\psi, v)$.*

Proof. Let $\psi(\cdot)$ be fixed, and $u_j(\cdot) \in L^2_{\mathbb{F}}[0,T]$ be a minimizing sequence of LQ problem, that is

$$\lim_{j\to\infty} J(\psi(\cdot), u_j(\cdot)) = \inf_{v(\cdot)\in\mathcal{U}} J(\psi(\cdot), v(\cdot)). \tag{26}$$

Let (Y^j, Z^j) be the state processes corresponding to $u_j(\cdot)$. It follows from (26) that there exists a constant M such that $J(\psi(\cdot), u_j(\cdot)) \leq M$ for any $j \geq 1$. Additionally, $J(\psi(\cdot), u_j(\cdot)) \geq \delta E\int_0^T |u_j(t)|^2 dt$, so we have $E\int_0^T |u_j(t)|^2 dt \leq \frac{M}{\delta}$. Consequently, there is a subsequence, which is still labeled by $u_j(\cdot)$, such that,

$$u_j(\cdot) \to u'(\cdot), \text{ weakly in } L^2_{\mathbb{F}}[0,T],$$

By Mazur's theorem, we have a sequence of convex combinations

$$\widehat{u}_j(\cdot) = \sum_{i\geq 1} \alpha_{ij} u_{i+j}(\cdot),\ \alpha_{i,j} \geq 0, \sum_{i\geq 1} \alpha_{ij} = 1,$$

such that

$$\widehat{u}_j(\cdot) \to u'(\cdot), \text{ strongly in } L^2_{\mathbb{F}}[0,T].$$

Since the set U is convex and closed, it follows that $u'(\cdot) \in \mathcal{U}$. On the other hand, the Theorem 3.7 in [26] leads to

$$E\int_0^T |\widehat{Y}_j(t) - Y'(t)|^2 dt \leq CE\int_0^T \left[\int_t^T (\widehat{u}_j(s) - u'(s))ds\right]^2 dt \leq CE\int_0^T |\widehat{u}_j(s) - u'(s)|^2 ds$$

i.e., $\widehat{Y}_j(\cdot) \to Y'(\cdot)$, strongly in $L^2_{\mathbb{F}}[0,T]$. By the convexity of the generator for (25),

$$\begin{aligned} J(\psi(\cdot), u'(\cdot)) &= \lim_{j\to\infty} J(\psi(\cdot), \widehat{u}_j(\cdot)) \\ &\leq \lim_{j\to\infty} \sum_{i\geq 1} \alpha_{ij} J(\psi(\cdot), u_{i+j}(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}} J(\psi(\cdot), u(\cdot)), \end{aligned}$$

which means that $u'(\cdot)$ is an optimal control. $\square$

In this setting, the maximum principle condition can be written as

$$\begin{aligned} &-R'(t)u^2(t) - l_2(0,t)u(t) - u(t)\int_0^t l_2(s,t)P(s)ds \\ &\geq -R'(t)u(t)v - l_2(0,t)v - v\int_0^t l_2(s,t)P(s)ds, \end{aligned} \tag{27}$$

with $v \in U$, and this leads to $R'(t)u(t) + l_2(0,t) + \int_0^t l_2(s,t)P(s)ds = 0$, i.e., $u(t) = R'^{-1}(t)[l_2(0,t) + \int_0^t l_2(s,t)P(s)ds]$, where

$$P(t) = Q'(t) + [l_1(0,t) + l_3(0,t)]G'Y(0) + \int_0^t l_1(s,t)P(s)ds + \int_0^t l_3(s,t)P(s)dW(s). \tag{28}$$

Hence $u(t)$ is the only control which satisfies the necessary conditions of optimality. By Lemma 4.3, it must be the unique optimal control. So we have

Theorem 4.3 *Let (H4) hold, there is a unique optimal control $u(\cdot)$ for the linear-quadratic control problem. Moreover, u has the representation: $u(t) = R'^{-1}(t)[l_2(0,t) + \int_0^t l_2(s,t)P(s)ds]$, where $P(s)$ satisfies (28).*

By the form of the optimal control, we deduce that the optimal control indeed a linear state feedback of the entire past history of the state process $P(\cdot)$ instead of being a feedback of the current state, which is similar to the result for linear-quadratic control of BSDEs, see p.6-p.7 in [6]. Substituting the representation of $u(\cdot)$ into (25), together with equation (28), we get the following coupled FBSVIE

$$\begin{cases} P(t) = Q'(t) + [l_1(0,t) + l_3(0,t)]G'Y(0) + \displaystyle\int_0^t l_1(s,t)P(s)ds + \int_0^t l_3(s,t)P(s)dW(s), \\ Y(t) = \psi(t) + \displaystyle\int_t^T l_2(s,t)R'^{-1}(s)l_2(0,s)ds + \int_t^T l_1(s,t)Y(s)ds \\ \qquad + \displaystyle\int_t^T l_2(s,t)\int_0^s l_2(u,s)P(u)duds + \int_t^T l_3(s,t)Z(s,t)ds - \int_t^T Z(t,s)dW(s), \end{cases} \tag{29}$$

Given (H4), by the unique existence of optimal control $u(\cdot)$, we observe that FBSVIE (29) admits a unique M-solution $(Y(\cdot), Z(\cdot,\cdot), P(\cdot))$ by which means that $P(\cdot)$ solves the forward equation of (29) in the Itô sense and $(Y(\cdot), Z(\cdot,\cdot))$ is the unique M-solution of the backward equation in (29). Consequently,

Theorem 4.4 *Let (H4) hold, then FBSVIE (29) admits a unique M-solution.*

5 Application in fincance

In this section, we illustrate the maximum principle above by studying the risk minimization problem in finance. Such kind of problem was studied by Mataramvura and Øksendal [9] by formulating it as a zero-sum stochastic differential game. Recently Øksendal and Sulem [12] also investigated this risk minimization problem via g-expectations. In this paper, we will consider the problem by means of the maximum principle in the previous. Some closed forms of the optimal solution are derived, which are consistent with the results in [12] or [23].

We consider a market with two investment possibilities, which are traded continuously until the fixed finite horizon T, is reached. One investment is described by

$$dS_0(t) = S_0(t)\rho(t)dt,\ S_0(0) = s_0.$$

The other financial instrument is described by

$$dS_1(t) = S_1(t)\alpha(t) + S_1(t)\beta(t)dW(t),\ S_1(0) = s_1.$$

Suppose the interest rate $\rho(\cdot)$ is nonnegative and bounded deterministic function, the stock-appreciation rate $\alpha(\cdot)$ and the stock-volatility $\beta(\cdot)$ are nonnegative and bounded

adapted processes. Moreover, $\beta^{-1}(\cdot)$ and $(\alpha(\cdot)-\rho(\cdot))^{-1}$ exist and bounded. The wealth process $X(\cdot)$ satisfies

$$dX(t) = [\rho(t)X(t) + v(t)(\alpha(t) - \rho(t))]dt + v(t)\beta(t)dW(t), \tag{30}$$

with $X(0) = x > 0$, thereby the solution of the wealth equation can be given by

$$X(t) = e^{\int_0^t \rho(s)ds}x + \int_0^t e^{\int_s^t \rho(u)du}[v(s)(\alpha(s) - \rho(s))ds + v(s)\beta(s)dW(s)].$$

A portfolio $v(\cdot)$, representing the amount invested in the risk asset, is said to be admissible if $v(\cdot) \in \mathcal{U}$. In the following the BSVIE that we are going to investigate is

$$Y(t) = -\psi(t) + \int_t^T [r(s)Y(s) + k_1(t,s)Z(s,t) + k_2(t,s)]ds - \int_t^T Z(t,s)dW(s), \tag{31}$$

where

$$\psi(t) = h(X(T)) + \int_t^T [l_1(t,s)X(s) + l_2(t,s)v(s)]ds.$$

Here r, l_i are bounded deterministic functions, k_i is a process such that $s \to k_i(t,s)$ is $\mathbb{F}$-progressively measurable for almost $t \in [0,T]$, k_1 is bounded, k_2 has the same assumption with g_0 in (H1). h is utility function of the terminal wealth that satisfies (H3) and $Eh^2(X(T)) < \infty$ (see the special cases below). If we define $\varrho(t;\psi(\cdot)) = Y(t)$, then by Theorem 3.6 in Yong [25], ϱ is a dynamic coherent risk measure. In order to show the result in a more explicit way, we would like to consider the special static case of $\varrho(0,\psi(\cdot)) = Y(0)$, and denote the cost functional by

$$\begin{aligned} J(v(\cdot)) &= Y(0) = -E\psi(0) + E\int_0^T [r(s)Y(s) + k_1(0,s)Z(s,0) + k_2(0,s)]ds \\ &= -Eh(X(T)) - E\int_0^T [l_1(0,s)X(s) + l_2(0,s)v(s)]ds \\ &\quad +E\int_0^T [r(s)Y(s) + k_1(0,s)EY(s) + k_2(0,s)]ds. \end{aligned}$$

We want to find $u \in \mathcal{U}$ such that $J(u(\cdot)) = \inf\limits_{v\in\mathcal{U}} J(v(\cdot))$. With the notation in the previous, we obtain $\gamma(y) = 0$, and $l(s,x,y,v) = [r(s)+Ek_1(0,s)]y - l_1(0,s)x - l_2(0,s)v + k_2(0,s)$. As to the coefficients in both (30) and (31),

$$\begin{aligned} b_x(t,s,x,v) &= \rho(s), \sigma_x(t,s,x,v) = 0, b_v(t,s,x,v) = \alpha(s) - \rho(s), \\ \sigma_v(t,s,x,v) &= \beta(s),\ g'_x(t,s) = l_1(t,s), \\ g'_v(t,s) &= l_2(t,s), g'_y(t,s) = r(s), g'_z(t,s) = k_1(t,s), \end{aligned}$$

where

$$g'_i(t,s) = g_i(t,s,x,y,z,v),\ i = x,v,y,z.$$

Then the Hamilton function is the form of

$$\begin{aligned}
& \mathcal{H}(t, X^u(t), Y^u(t), Z^u(t,\cdot), u(t), P(t), Q(t), R(\cdot,t), v) \\
= & -v\left[(\alpha(t)-\rho(t))E^{\mathcal{F}_t}h_x(X^u(T)) + \beta(t)\pi(t)\right] \\
& -v\left[\int_0^t l_2(s,t)P(s)ds + \beta(t)E^{\mathcal{F}_t}\int_t^T R(s,t)ds + (\alpha(t)-\rho(t))E^{\mathcal{F}_t}\int_t^T Q(s)ds\right],
\end{aligned}$$

and the adjoint equation

$$\begin{cases}
P(t) = [r(t) + Ek_1(0,t)] + r(t)\int_0^t P(s)ds + \int_0^t k_1(s,t)P(s)dW(s), \\
Q(t) = -l_1(0,t) - \rho(t)h_x(X^u(T)) + \int_0^t l_1(s,t)P(s)ds \\
\qquad + \int_t^T \rho(t)Q(s)ds - \int_t^T R(t,s)dW(s), \\
h_x(X^u(T)) = Eh_x(X^u(T)) + \int_0^T \pi(s)dW(s).
\end{cases} \tag{32}$$

Since $\mathcal{H}$ is a linear function in v, then the coefficient of v vanishes, i.e.,

$$0 = (\alpha(t)-\rho(t))E^{\mathcal{F}_t}h_x(X^u(T)) + \beta(t)\pi(t) + M(t), \tag{33}$$

where

$$M(t) = \int_0^t l_2(s,t)P(s)ds + \beta(t)E^{\mathcal{F}_t}\int_t^T R(s,t)ds + (\alpha(t)-\rho(t))E^{\mathcal{F}_t}\int_t^T Q(s)ds. \tag{34}$$

The lemma below is essentially similar to Theorem A.2 in the appendix of [12]. For readers' convenience, we present a proof here.

Lemma 5.1 *Consider the equation, $\alpha_1(t)E^{\mathcal{F}_t}\xi + \beta_1(t)\theta(t) = \zeta(t)$, where ξ is a $\mathcal{F}_T$-measurable random variable satisfying $\xi = E\xi + \int_0^T \theta(s)dW(s)$. Assume $\beta_1^{-1}(t)$ exists, $\zeta(t)$ is a adapted process, then ξ must be the form of*

$$\xi = e^{-A_1(T)}E\xi + e^{-A_1(T)}\int_0^T e^{A_1(s)}\beta_1^{-1}(s)\zeta(s)dW(s),$$

where A_1 is given by

$$A_1(t) = \int_0^t \beta_1^{-1}(s)\alpha_1(s)dW(s) + \frac{1}{2}\int_0^t \beta_1^{-2}(s)\alpha_1^2(s)ds. \tag{35}$$

Proof. We denote by $P(t) = E^{\mathcal{F}_t}\xi$, therefore $P(t) = P(0) + \int_0^t \theta(s)dW(s)$. On the other hand,

$$\theta(t) = \beta_1^{-1}(t)[\zeta(t) - \alpha_1(t)P(t)],$$

so

$$P(t) = P(0) + \int_0^t \beta_1^{-1}(s)[\zeta(s) - \alpha_1(s)P(s)]dW(s).$$

Since we can rewrite P by

$$P(t) = e^{-A_1(t)}P(0) + e^{-A_1(t)}\int_0^t e^{A_1(s)}\beta_1^{-1}(s)\zeta(s)dW(s),$$

with A_1 given by (35), thereby

$$\xi = e^{-A_1(T)}E\xi + e^{-A_1(T)}\int_0^T e^{A_1(s)}\beta_1^{-1}(s)\zeta(s)dW(s).$$

The conclusion follows clearly. $\square$

Remark 5.1 *There are two things worthy to point out. On the one hand, if we assume that $h_x(X^u(T)) \in \mathbb{D}_{1,2}$, see [11], then by Ocone-Clark formula, $\pi(t) = E^{\mathcal{F}_t}D_t h_x(X^u(T))$, then (33) can be rewritten as*

$$(\alpha(t) - \rho(t))E^{\mathcal{F}_t}h_x(X^u(T)) + \beta(t)E^{\mathcal{F}_t}D_t h_x(X^u(T)) + M(t) = 0.$$

It is a linear inhomogeneous Malliavin-differential type equation in the unknown random variable $h_x(X^u(T))$, which can also seen in [10]and [12]. On the other hand, if $\zeta = 0$, there are infinite random variables satisfying the equation in Lemma 5.1. For example, if ξ_1 satisfies it, so does $c\xi_1$, with c being a constant.

To sum up, we give

Theorem 5.1 *Suppose u is an optimal portfolio of the above risk minimizing problem, then $u(\cdot)$ must satisfies*

$$h_x(X^u(T)) = e^{-A(T)}Eh_x(X^u(T)) - e^{-A(T)}\int_0^T e^{A(s)}\beta^{-1}(s)M(s)dW(s),$$

with $X^u(T)$ be the terminal wealth corresponding to u, $M(t)$ given by (34), and

$$A(t) = \int_0^t \beta^{-1}(s)(\alpha(s) - \rho(s))dW(s) + \frac{1}{2}\int_0^t \beta^{-2}(s)(\alpha(s) - \rho(s))^2 ds.$$

In order to express the explicit form of u, some more assumptions are required. Let $l_i = 0$, $(i = 1, 2)$, then the Hamilton function becomes

$$\begin{aligned}\mathcal{H} &= -v[(\alpha(t) - \rho(t))E^{\mathcal{F}_t}h_x(X^u(T)) + \beta(t)\pi(t)] \\ &\quad -v\left(\beta(t)E^{\mathcal{F}_t}\int_t^T R(s,t)ds + (\alpha(t) - \rho(t))E^{\mathcal{F}_t}\int_t^T Q(s)ds\right),\end{aligned}$$

where

$$\begin{cases} Q(t) = -\rho(t)h_x(X^u(T)) + \int_t^T \rho(t)Q(s)ds - \int_t^T R(t,s)dW(s), \\ h_x(X^u(T)) = Eh_x(X^u(T)) + \int_0^T \pi(s)dW(s), \end{cases}$$

therefore, the optimal portfolio u satisfies

$$\begin{aligned} 0 &= (\alpha(t)-\rho(t))E^{\mathcal{F}_t}h_x(X^u(T)) + \beta(t)\pi(t) \\ &\quad +\beta(t)E^{\mathcal{F}_t}\int_t^T R(s,t)ds + (\alpha(t)-\rho(t))E^{\mathcal{F}_t}\int_t^T Q(s)ds. \end{aligned} \tag{36}$$

As a consequence, by solving the above simple BSVIE, we deduce that $\forall (t,s)\in\Delta$, i.e., $0\le s<t\le T$,

$$Q(t) = -\rho(t)e^{\int_t^T \rho(u)du}\cdot E^{\mathcal{F}_t}h_x(X^u(T)), R(t,s) = -\rho(t)e^{\int_t^T \rho(u)du}\pi(s). \tag{37}$$

Substituting (37) into (36), we arrive at

$$(\alpha(t)-\rho(t))E^{\mathcal{F}_t}h_x(X(T)) + \beta(t)\pi(t) = 0. \tag{38}$$

Recalling Lemma 5.1, we can express $h_x(X^u(T))$ as $h_x(X^u(T)) = e^{-A(T)}Eh_x(X^u(T))$, which is a necessary condition for u being optimal.

Remark 5.2 *On the one hand, due to (38) being a homogeneous Malliavin-differential type equation, if $h(\cdot)$ is replaced with $ch(\cdot)$ in the cost functional, with c being a constant, we can still obtain the same result. On the other hand, thanks to $l_2 = 0$, the Hamilton function $\mathcal{H}$ is independent of P, which is solution of the forward adjoint equation in (32). Then no matter what values of k_1, it does not change the value of optimal portfolio u.*

Now we will prove that the necessary condition above is also sufficient. For any $v_i\in\mathcal{U}$ with $i=1,2$, we have form the concavity of h that $Eh(X^{v_1}(T)) - Eh(X^{v_2}(T)) \ge E[h_x(X^{v_1}(T))(X^{v_1}(T)-X^{v_2}(T))]$, where $X^{v_i}(T)$ is the terminal wealth corresponding to v_i, thus one sufficient condition for the strategy u being optimal is that $E[h_x(X^u(T))X^v(T)]$ being a constant over $v\in\mathcal{U}$, see Proposition 2.1 in Z. Wang [23]. By the above necessary condition we have $Eh_x(X^u(T))X^v(T) = Eh_x(X^u(T))\cdot Ee^{-A(T)}X^v(T)$. Using Itô formula to $e^{-A(t)}X^v(t)e^{\int_t^T \rho(s)ds}$, one gets

$$\begin{aligned} &e^{-A(T)}X^v(T) - xe^{\int_0^T \rho(s)ds} \\ = &\int_0^T e^{-A(s)}e^{\int_s^T \rho(s)ds}\left[v(s)\beta(s) - X(s)\frac{\alpha(s)}{\beta(s)}\right]dW(s), \end{aligned}$$

hence $Ee^{-A(T)}X^v(T) = xe^{\int_0^T \rho(s)ds}$, $Eh_x(X^u(T))X^v(T)$ is a constant independent of v. Thus we obtain

Theorem 5.2 *Suppose u is an optimal portfolio of the above risk minimizing problem if and only if $u(\cdot)$ satisfies $h_x(X^u(T)) = e^{-A(T)}Eh_x(X^u(T))$.*

Remark 5.3 *There is one thing worthy to point out. The above argument also implies that the sufficient condition in Proposition 2.1 in [23] is also necessary. In fact, when h is a concave utility function, from our stochastic maximum principle we know that one necessary condition for u being optimal is equation (38), and this also implies that $Ee^{-A(T)}X^v(T)$ is a constant independent of v.*

Some special cases for the function $h(x)$ are given below to show some exact results of u.

Case 1 $h(x) = x$, we will deduce that the optimal portfolio is $u(t) = 0$. In fact, from equation (38), we have $\alpha(t) - \rho(t) = 0$, then $\alpha(t) = \rho(t)$, which means the stock-appreciation rate is equal to the interest rate. In this case, the optimal portfolio is $u(t) = 0$. As we know, there are usually many kinds of method, i.e., risk measure, to measure the wealth at some time, for example the terminal wealth at time T, and there is one optimal portfolio for each kind of risk measure. On the other hand, from Remark 5.2, we can choose any bounded $k_1(t,s)$, in other words different risk measures, thus get different value $Y(0)$, i.e., different minimal risk, while the optimal portfolio is the same one. For example, if $k_1(t,s) = 0$, $k_2(t,s)$ is independent of t, then the minimal risk can be expressed as, $Y(0) = e^{\int_0^T \rho(s)+r(s)ds}x + E\int_0^T e^{\int_0^s r(u)du}k_2(s)ds$. Note that $\rho = 0$, it is consistent with the one in [12].

Remark 5.4 *Recently, the author [20] consider the case when r is allowed to be random, while r is assumed to be deterministic in [25]. In this case, it is easy to show that the above results also hold, and the minimal risk is given by $Y(0) = Ee^{\int_0^T \rho(s)+r(s)ds}x + E\int_0^T e^{\int_0^s r(u)du}k_2(s)ds$.*

Case 2 $h(x) = x - \frac{\gamma}{2}x^2$ with $\gamma \neq 0$ being a constant, then $h_x(x) = 1 - \gamma x$, and in the following, we denote $F(t) = E^{\mathcal{F}_t}h_x(X^u(T)) = 1 - \gamma E^{\mathcal{F}_t}X(T)$. In this setting, we assume that $E\int_0^T |v(s)|^4 ds < \infty$. Following the idea in [23], we will show the explicit form of the optimal portfolio u. By equation (38), we know that $\pi(t) = -\beta^{-1}(t)F(t)(\alpha(t) - \rho(t))$. Using Itô formula to $A(t)F(t)$ on $[0,T]$, where $A(t) = e^{\int_0^t a(s)ds}$ and a is a deterministic integral function,

$$A(T)F(T) = F(0) + \int_0^T A(s)\pi(s)dW(s) + \int_0^T F(s)a(s)A(s)ds. \tag{39}$$

Since $F(T) = 1 - \gamma X^u(T)$, together with (39), we have

$$\begin{aligned} X^u(T) &= \gamma^{-1} - (\gamma A(T))^{-1}\left[F(0) + \int_0^T A(s)\pi(s)dW(s) + \int_0^T F(s)a(s)A(s)ds\right] \\ &= \gamma^{-1} - \frac{F(0)}{\gamma A(T)} - \int_0^T \frac{F(s)a(s)A(s)}{\gamma A(T)}ds + \int_0^T \frac{A(s)F(s)(\alpha(s)-\rho(s))}{\gamma A(T)\beta(s)}dW(s). \end{aligned}$$

On the other hand, by (30),

$$X^u(T) = e^{\int_0^T \rho(s)ds}x + \int_0^T \left[e^{\int_s^T \rho(u)du}(\alpha(s)-\rho(s))u(s)ds + e^{\int_s^T \rho(u)du}\beta(s)u(s)dW(s)\right], \quad (40)$$

then by comparing the correspondent part in (39) and (40),

$$\begin{cases} e^{\int_0^T \rho(s)ds} x = \gamma^{-1} - \frac{F(0)}{\gamma A(T)}, \\ -\frac{F(s)a(s)A(s)}{\gamma A(T)} = e^{\int_s^T \rho(u)du}(\alpha(s)-\rho(s))u(s), \\ \frac{A(s)F(s)(\alpha(s)-\rho(s))}{\gamma A(T)\beta(s)} = e^{\int_s^T \rho(u)du}\beta(s)u(s), \end{cases}$$

thereby we deduce that $a(s) = \frac{\beta^2(s)}{|\alpha(s)-\rho(s)|^2}$, and the optimal portfolio is expressed as

$$\begin{aligned} u(s) &= e^{-\int_s^T \rho(u)du}\frac{(\alpha(s)-\rho(s))F(s)A(s)}{\beta^2(s)\gamma A(T)} \\ &= e^{-\int_s^T (\frac{|\alpha(u)-\rho(u)|^2}{\beta^2(u)}+\rho(u))du}\frac{(\alpha(s)-\rho(s))}{\beta^2(s)}\left[\frac{1}{\gamma} - E^{\mathcal{F}_s}X^u(T)\right]. \end{aligned} \quad (41)$$

To get more feeling about the general form of equation (33) in the previous, i.e., the inhomogeneous Malliavin differential equation, we will consider some special cases. Let $l_i(0,t) = l_i(t)$, $i = 1, 2$. Moreover, for the sake of obtaining the exact expression of P, Q, R, we assume $k_1(t,s) = r(s)$.

Case 1 If $\rho(t) = 0$, then for any $t \in [0,T]$, $s < t$,

$$\begin{aligned} P(t) &= 2r(t)e^{\int_0^t r(s)ds - \frac{1}{2}\int_0^t r^2(s)ds + \int_0^t r(s)dW(s)}, \\ Q(t) &= -l_1(t) + l_1(t)\int_0^t P(s)ds \\ &= l_1(t)\left(-1 + 2\int_0^t r(s)e^{\int_0^s r(u)du - \frac{1}{2}\int_0^s r^2(u)du + \int_0^s r(u)dW(u)}ds\right), \\ R(t,s) &= 2l_1(t)\left(e^{\int_0^t r(u)du} - e^{\int_0^s r(u)du}\right)r(s)e^{-\frac{1}{2}\int_0^s r^2(u)du + \int_0^s r(u)dW(u)}, \end{aligned}$$

then

$$\begin{aligned} M(t) &= 2l_2(t)\int_0^t r(s)e^{\int_0^s r(u)du - \frac{1}{2}\int_0^s r^2(u)du + \int_0^s r(u)dW(u)}ds \\ &\quad + 2\beta(t)E^{\mathcal{F}_t}\int_t^T l_1(s)\left(e^{\int_0^s r(u)du} - e^{\int_0^t r(u)du}\right)r(t)e^{-\frac{1}{2}\int_0^t r^2(u)du + \int_0^t r(u)dW(u)}ds \\ &\quad + \alpha(t)E^{\mathcal{F}_t}\int_t^T l_1(s)\left(-1 + 2\int_0^s r(u)e^{\int_0^u r(v)dv - \frac{1}{2}\int_0^u r^2(v)dv + \int_0^u r(v)dW(v)}du\right)ds. \end{aligned}$$

In fact, in this case, $l_y(s,x,y,v)=2r(s)$, and $P(\cdot)$ and $Q(\cdot)$ satisfy

$$\begin{cases} P(t)=2r(t)+r(t)\displaystyle\int_0^t P(s)ds+r(t)\int_0^t P(s)dW(s), \\ Q(t)=-l_1(t)+l_1(t)\displaystyle\int_0^t P(s)ds-\int_t^T R(t,s)dW(s). \end{cases}$$

By the martingale representation theorem, there exists a unique adapted process such that

$$P(t)=EP(t)+\int_0^t T(t,s)dW(s),$$

thus we have

$$R(t,s)=l_1(t)\int_s^t T(u,s)du.$$

Assume $r\neq 0$, and $p'(t)=\frac{P(t)}{r(t)}$, then we have

$$P'(t)=2+\int_0^t r(s)P'(s)ds+\int_0^t r(s)P'(s)dW(s),$$

Using Ito formula to $P''(t)=e^{-\int_0^t r(u)du}P'(t)$ and we obtain $P''(t)=2+\int_0^t r(s)P''(s)dW(s)$, thus we solve that $P''(t)=2e^{-\frac{1}{2}\int_0^t r^2(s)ds+\int_0^t r(s)dW(s)}$, so we obtain $P(t)$ above. On the other hand, $EP''(t)=2$, then $P''(t)=EP''(t)+\int_0^t r(s)P''(s)dW(s)$, thus

$$e^{-\int_0^t r(u)du}\frac{P(t)}{r(t)}=e^{-\int_0^t r(u)du}\frac{EP(t)}{r(t)}+\int_0^t r(s)P''(s)dW(s),$$

thus

$$T(t,s)=2r(t)r(s)e^{\int_0^t r(u)du-\frac{1}{2}\int_0^s r^2(u)du+\int_0^s r(u)dW(u)},$$

thus we get

$$R(t,s)=2l_1(t)r(s)e^{-\frac{1}{2}\int_0^s r^2(u)du+\int_0^s r(u)dW(u)}\left[e^{\int_0^t r(v)dv}-e^{\int_0^s r(v)dv}\right].$$

As to the case $r=0$, then $\forall(t,s)\in\Delta$, $P(t)=0$, $Q(t)=-l_1(t)$, $R(t,s)=0$, and they are all consistent with the above results.

Case 2 If $r(s)=0$, then we have $P(t)=0$, and

$$Q(t)=-l_1(t)-\rho(t)h_x(X^u(T))+\int_t^T \rho(t)Q(s)ds-\int_t^T R(t,s)dW(s).$$

and we can solve M-solution as follows

$$\begin{aligned} Q(t) &= -l_1(t)-\rho(t)e^{\int_t^T \rho(u)du}\cdot E^{\mathcal{F}_t}h_x(X^u(T)) \\ &\quad -\rho(t)e^{\int_t^T \rho(u)du}\cdot\int_t^T e^{-\int_s^T \rho(u)du}l_1(s)ds, \end{aligned}$$

and

$$R(t,s) = -\rho(t)e^{\int_t^T \rho(s)ds}\pi(s), \quad (t,s) \in \Delta,$$

thus

$$\begin{aligned} M(t) &= \big(\beta(t)\pi(t) + (\alpha(t)-\rho(t))E^{\mathcal{F}_t}h_x(X^u(T))\big)\left(1 - e^{\int_t^T \rho(u)du}\right) \\ &\quad -(\alpha(t)-\rho(t))E^{\mathcal{F}_t}\int_t^T \left(l_1(s) + \rho(s)\int_s^T e^{\int_s^v \rho(u)du} l_1(v)dv\right) ds. \end{aligned}$$

Case 3 $r(s) \neq 0$, $\rho(s) \neq 0$. In this case, we can obtain the following result by combining the results in above two cases together,

$$\begin{aligned} P(t) &= 2r(t)e^{\int_0^t r(s)ds - \frac{1}{2}\int_0^t r^2(s)ds + \int_0^t r(s)dW(s)}, \\ Q(t) &= Q'(t) - \rho(t)e^{\int_t^T \rho(u)du}\cdot E^{\mathcal{F}_t}h_x(X(T)) + \rho(t)E^{\mathcal{F}_t}\int_t^T e^{\int_t^s \rho(u)du} Q'(s)ds, \\ R(t,s) &= R'(t,s) - \rho(t)e^{\int_t^T \rho(u)du}\pi(s) + \rho(t)\int_t^T e^{\int_t^u \rho(v)dv} R'(u,s)du, \end{aligned}$$

where

$$\begin{aligned} Q'(t) &= l_1(t)\left(-1 + 2\int_0^t r(s)e^{\int_0^s r(u)du - \frac{1}{2}\int_0^s r^2(u)du + \int_0^s r(u)dW(u)} ds\right) \\ R'(t,s) &= 2l_1(t)\left(e^{\int_0^t r(u)du} - e^{\int_0^s r(u)du}\right) r(s)e^{-\frac{1}{2}\int_0^s r^2(u)du + \int_0^s r(u)dW(u)}, \end{aligned}$$

thus

$$\begin{aligned} M(t) &= (\alpha(t)-\rho(t))E^{\mathcal{F}_t}h_x(X^u(T))\left(1 - e^{\int_t^T \rho(u)du}\right) + \beta(t)\pi(t)\left(1 - e^{\int_t^T \rho(u)du}\right) \\ &\quad +(\alpha(t)-\rho(t))E^{\mathcal{F}_t}\int_t^T \left(Q'(s) + \rho(s)\int_s^T e^{\int_s^v \rho(u)du} Q'(v)dv\right) ds \\ &\quad +l_2(t)\left(\int_0^t 2r(s)e^{\int_0^s r(u)du - \frac{1}{2}\int_0^s r^2(u)du + \int_0^s r(u)dW(u)} ds\right) \\ &\quad +\beta(t)E^{\mathcal{F}_t}\int_t^T \left(R'(s,t) + \rho(s)\int_s^T e^{\int_s^v \rho(u)du} R'(v,t)dv\right) ds. \end{aligned}$$

References

[1] V. Anh, J. Yong, *Backward stochastic Volterra integral equations in Hilbert space,* In: Differential and Difference Equations and Applications, Hindawi, New York (2006) pp. 57-66.

[2] S.A. Belbas, *Iterative schemes for optimal control of Volterra integral equations,* Nonlinear Anal. **37** (1999) 57-79.

[3] S.A. Belbas, *A new method for optimal control of Volterra integral equations,* Appl. Math. Comput. **189** (2007) 1902-1915.

[4] N. El Karoui, S. Peng and M. Quenez, *Backward stochastic differential equations in finance,* Math. Finance. **7** (1997) 1-71.

[5] I. Karatzas, S. E. Shreve, *Brownian Motion and Stochastic Calculus,* Springer, Heidelberg (1988).

[6] E. B. Lim, Y. Zhou, *Linear-quadratic control of backward stochastic differential equations,* SIAM J. Contr. Optim. **40** (2001) 450-474.

[7] J. Lin, *Adapted solution of backward stochastic nonlinear Volterra integral equation,* Stoch. Anal. Appl. **20** (2002) 165-183.

[8] J. Ma, J. Yong, *"Forward-Dackward Stochastic Differential Equations and Their Applications,"* Lecture Notes in Math. Vol. 1702, Springer-Verlag, Berlin, 1999.

[9] S. Mataramvura, B. Øksendal *Risk minimizing portfolio and HJBI equations for stochastic differential games,* Stochatics **80** (2008) 317-337.

[10] T. Meyer-Brandis, B. Øksendal, X. Y. Zhou, *A Malliavin calculus approach to a general maximum principle for stochastic control,* preprint.

[11] D. Nualart, *"The Malliavin Calculus and Related Topics,"* Probability and Its Applications, Springer-Verlag, New York and Berlin, 1995.

[12] B. Øksendal, A. Sulem, *Maximum principle for optimal control of forward-backward stochastic differential equations with jumps,* preprint.

[13] B. Øksendal, T. Zhang, *Optimal control with partial information for stochastic Volterra equations,* preprint.

[14] E. Pardoux, S. Peng, *Adapted solution of a backward stochastic differential equation,* Sys. Control Letters **14** (1990) 55-61.

[15] S. Peng, *A general stochastic maximum principle for optimal control problems,* SIAM J.Control Optim. **28** (1990) 966-979.

[16] S. Peng, *Backward stochastic differential equations and application to optimal control,* Appl. Math. Optim.**27** (1993) 125-144.

[17] A. J. Pritchard, Y. You, *Causal feedback optimal control for Volterra integral equations,* SIAM J. Control Optim. **34** (1996) 1874-1890.

[18] Y.Ren, *On solutions of Backward stochastic Volterra integral equations with jumps in hilbert spaces,* J Optim Theory Appl **144** (2010) 319-333.

[19] V. R. Vinokurov, *Optimal control of processes described by integral equations,* SIAM J. Control **7** (1969) 324-355.

[20] T. Wang, *BSVIEs with stochastic Lipschitz coefficients and applications in finance,* preprint(pdf-file available in arXiv: math. PR/1001.3558v1 20 Jan 2010)

[21] Z. Wang, X. Zhang, *Non-Lipschitz backward stochastic volterra type equations with jumps,* Stoch.Dyn. **7** (2007), 479-496.

[22] T. Wang, Y. Shi, *Solvabiltiy of general backward stochastic Volterra integral equation with non-Lipschitz conditions,* preprint(pdf-file available in arXiv: math. PR/1001.3557v1 20 Jan 2010)

[23] Z. Eang, J. Xia, L. Zhang, *Optimal investment for an insurer: The martingale approach,* Insurance: Math. Econ. **40** (2007), 322-334.

[24] J. Yong, *Backward stochastic Volterra integral equations and some related problems,* Stochastic Proc. Appl. **116** (2006) 779–795.

[25] J. Yong, *Continuous-time dynamic risk measures by backward stochastic Volterra integral equations,* Appl. Anal. **86** (2007) 1429–1442.

[26] J. Yong, *Well-posedness and regularity of backward stochastic Volterra integral equation,* Probab. Theory Relat. Fields. **142** (2008) 21-77.

[27] Y. You, *Quadratic integral games and causal synthesis,* Trans. Amer. Math. Soc. **352** (2000) 2737-2764.

[28] J. Yong, X.Y. Zhou, *Stochstic Control: Hamiltonian Systems and HJB Equations,* Springer-Verlag, New York, 1999